# Frequently Used Statistics Formulas and Tables

## Chapter 2

Class Width = $\frac{\text{highest value} - \text{lowest value}}{\text{number classes}}$ (increase to next integer)

Class Midpoint = $\frac{\text{upper limit} + \text{lower limit}}{2}$

## Chapter 3

$n$ = sample size

$N$ = population size

$f$ = frequency

$\sum$ = sum

$w$ = weight

Sample mean: $\bar{x} = \frac{\sum x}{n}$

Population mean: $\mu = \frac{\sum x}{N}$

Weighted mean: $\bar{x} = \frac{\sum (w \cdot x)}{\sum w}$

Mean for frequency table: $\bar{x} = \frac{\sum (f \cdot x)}{\sum f}$

Midrange = $\frac{\text{highest value} + \text{lowest value}}{2}$

Range = Highest value - Lowest value

Sample standard deviation: $s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$

Population standard deviation: $\sigma = \sqrt{\frac{\sum (x - \mu)^2}{N}}$

Sample variance: $s^2$

Population variance: $\sigma^2$

## Chapter 3

Limits for Unusual Data

Below: $\mu - 2\sigma$

Above: $\mu + 2\sigma$

Empirical Rule

About 68%: $\mu - \sigma$ to $\mu + \sigma$

About 95%: $\mu - 2\sigma$ to $\mu + 2\sigma$

About 99.7%: $\mu - 3\sigma$ to $\mu + 3\sigma$

Sample coefficient of variation: $CV = \frac{s}{\bar{x}} \cdot 100\%$

Population coefficient of variation: $CV = \frac{\sigma}{\mu} \cdot 100\%$

Sample standard deviation for frequency table: $s = \sqrt{\frac{n[\sum (f \cdot x^2)] - [\sum (f \cdot x)]^2}{n(n-1)}}$

Sample z-score: $z = \frac{x - \bar{x}}{s}$

Population z-score: $z = \frac{x - \mu}{\sigma}$

Interquartile Range: (IQR) = $Q3 - Q1$

Modified Box Plot Outliers

lower limit: Q1 - 1.5 (IQR)

upper limit: Q3 + 1.5 (IQR)

Chapter 4

Probability of the complement of event *A*
$P(\text{not } A) = 1 - P(A)$

Multiplication rule for independent events $P(A \text{ and } B) = P(A) \cdot P(B)$

General multiplication rules
$P(A \text{ and } B) = P(A) \cdot P(B, \text{ given } A)$
$P(A \text{ and } B) = P(A) \cdot P(A, \text{ given } B)$

Addition rule for mutually exclusive events
$P(A \text{ or } B) = P(A) + P(B)$

General addition rule
$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

Permutation rule: $_nP_r = \frac{n!}{(n-r)!}$

Combination rule: $_nC_r = \frac{n!}{r!(n-r)!}$

Permutation and Combination on TI 83/84

*n* Math PRB nPr enter *r*

*n* Math PRB nCr enter *r*

Note: textbooks and formula sheets interchange "r" and "x" for number of successes

Chapter 5

Discrete Probability Distributions:

Mean of a discrete probability distribution:
$\mu = \sum[x \cdot P(x)]$

Standard deviation of a probability distribution:

$\sigma = \sqrt{\sum[x^2 \cdot P(x)] - \mu^2}$

Binomial Distributions
$r$ = number of successes (or x)
$p$ = probability of success
$q$ = probability of failure
$q = 1 - p$  $p + q = 1$

Binomial probability distribution
$P(r) = C_{n,r}\, p^r q^{n-r}$
Mean: $\mu = np$

Standard deviation: $\sigma = \sqrt{npq}$

Poisson Distributions

$r$ = number of successes (or $x$)

$\mu$ = mean number of successes (over a given interval)
Poisson probability distribution

$P(r) = \frac{e^{-\mu}\mu^r}{r!}$

$e \approx 2.71828$

$\mu$ = mean (over some interval)

$\sigma = \sqrt{\mu}$

$\sigma^2 = \mu$

Chapter 6

Normal Distributions

Raw score: $x = z\sigma + \mu$

Standard score: $z = \frac{x - \mu}{\sigma}$

Mean of $\bar{x}$ distribution: $\mu_{\bar{x}} = \mu$

Standard deviation of $\bar{x}$ distribtuion: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$
(standard error)

Standard score for $\bar{x}$: $z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$

Chapter 7

One Sample Confidence Interval

*for proportions ($p$): ( $np > 5$ and $nq > 5$)

$\hat{p} - E < p < \hat{p} + E$

where $E = z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$\hat{p} = \frac{r}{n}$

for means ($\mu$) when $\sigma$ is known:

$\bar{x} - E < \mu < \bar{x} + E$

where $E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

for means ($\mu$) when $\sigma$ is unknown:

$\bar{x} - E < \mu < \bar{x} + E$

where $E = t_{\alpha/2}\frac{s}{\sqrt{n}}$

with $d.f. = n - 1$

for variance ($\sigma^2$): $\frac{(n-1)s^2}{\chi^2_R} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_L}$

with $d.f. = n - 1$

Chapter 7

Confidence Interval: Point estimate ± error

Point estimate = $\frac{\text{Upper limit + Lower limit}}{2}$

Error = $\frac{\text{Upper limit - Lower limit}}{2}$

Sample Size for Estimating

means:

$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2$

proportions:

$n = \hat{p}\hat{q}\left(\frac{z_{\alpha/2}}{E}\right)^2$ with preliminary estimate for $p$

$n = 0.25\left(\frac{z_{\alpha/2}}{E}\right)^2$ without preliminary estimate for $p$

variance or standard deviation:
*see table 7-2 (last page of formula sheet)

Confidence Intervals

| Level of Confidence | z-value ($z_{\alpha/2}$) |
|---|---|
| 70% | 1.04 |
| 75% | 1.15 |
| 80% | 1.28 |
| 85% | 1.44 |
| 90% | 1.645 |
| 95% | 1.96 |
| 98% | 2.33 |
| 99% | 2.58 |

## Chapter 8

### One Sample Hypothesis Testing

for $p$ ($np>5$ and $nq>5$): $z = \dfrac{\hat{p}-p}{\sqrt{pq/n}}$

where $q = 1-p$; $\hat{p} = r/n$

for $\mu$ ($\sigma$ known): $z = \dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}$

for $\mu$ ($\sigma$ unknown): $t = \dfrac{\bar{x}-\mu}{s/\sqrt{n}}$ with $d.f. = n-1$

for $\sigma^2$: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2}$ with $d.f. = n-1$

## Chapter 9

### Two Sample Confidence Intervals and Tests of Hypotheses

### Difference of Proportions ($p_1 - p_2$)

Confidence Interval:

$(\hat{p}_1-\hat{p}_2)-E < (p_1-p_2) < (\hat{p}_1-\hat{p}_2)+E$

where $E = z_{\alpha/2}\sqrt{\dfrac{\hat{p}_1\hat{q}_1}{n_1}+\dfrac{\hat{p}_2\hat{q}_2}{n_2}}$

$\hat{p}_1 = r_1/n_1$; $\hat{p}_2 = r_2/n_2$ and $\hat{q}_1 = 1-\hat{p}_1$; $\hat{q}_2 = 1-\hat{p}_2$

Hypothesis Test:

$z = \dfrac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\dfrac{\bar{p}\bar{q}}{n_1}+\dfrac{\bar{p}\bar{q}}{n_2}}}$

where the pooled proportion is $\bar{p}$

$\bar{p} = \dfrac{r_1+r_2}{n_1+n_2}$ and $\bar{q} = 1-\bar{p}$

$\hat{p}_1 = r_1/n_1$; $\hat{p}_2 = r_2/n_2$

## Chapter 9

### Difference of means $\mu_1-\mu_2$ (independent samples)

Confidence Interval when $\sigma_1$ and $\sigma_2$ are known

$(\bar{x}_1-\bar{x}_2)-E < (\mu_1-\mu_2) < (\bar{x}_1-\bar{x}_2)+E$

where $E = z_{\alpha/2}\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}$

Hypothesis Test when $\sigma_1$ and $\sigma_2$ are known

$z = \dfrac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$

Confidence Interval when $\sigma_1$ and $\sigma_2$ are unknown

$(\bar{x}_1-\bar{x}_2)-E < (\mu_1-\mu_2) < (\bar{x}_1-\bar{x}_2)+E$

$E = t_{\alpha/2}\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}$

with $d.f.$ = smaller of $n_1-1$ and $n_2-1$

Hypothesis Test when $\sigma_1$ and $\sigma_2$ are unknown

$t = \dfrac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}}$

with $d.f.$ = smaller of $n_1-1$ and $n_2-1$

### Matched pairs (dependent samples)

Confidence Interval

$\bar{d}-E < \mu_d < \bar{d}+E$

where $E = t_{\alpha/2}\dfrac{s_d}{\sqrt{n}}$ with d.f. = $n-1$

Hypothesis Test

$t = \dfrac{\bar{d}-\mu_d}{s_d/\sqrt{n}}$ with $d.f. = n-1$

### Two Sample Variances

Confidence Interval for $\sigma_1^2$ and $\sigma_2^2$

$\left(\dfrac{s_1^2}{s_2^2}\cdot\dfrac{1}{F_{right}}\right) < \dfrac{\sigma_1^2}{\sigma_2^2} < \left(\dfrac{s_1^2}{s_2^2}\cdot\dfrac{1}{F_{left}}\right)$

Hypothesis Test Statistic: $F = \dfrac{s_1^2}{s_2^2}$ where $s_1^2 \geq s_2^2$

numerator $d.f. = n_1-1$ and denominator $d.f. = n_2-1$

.

## Chapter 10

## Regression and Correlation

Linear Correlation Coefficient (r)

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2}\sqrt{n(\sum y^2) - (\sum y)^2}}$$

OR

$$r = \frac{\sum (z_x z_y)}{n-1}$$ where $z_x$ = z score for x and $z_y$ = z score for y

Coefficient of Determination: $r^2 = \frac{\text{explained variation}}{\text{total variation}}$

Standard Error of Estimate: $s_e = \sqrt{\frac{\sum (y-\hat{y})^2}{n-2}}$ or $s_e = \sqrt{\frac{\sum y^2 - b_0 \sum y - b_1 \sum xy}{n-2}}$

Prediction Interval: $\hat{y} - E < y < \hat{y} + E$

where $E = t_{\alpha/2}\, s_e \sqrt{1 + \frac{1}{n} + \frac{n(x_0 - \bar{x})^2}{n(\sum x^2) - (\sum x)^2}}$

Sample test statistic for $r$

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} \text{ with } d.f. = n-2$$

Least-Squares Line (Regression Line or Line of Best Fit)

$\hat{y} = b_0 + b_1 x$ note that $b_0$ is the y-intercept and $b_1$ is the slope

where $b_1 = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$ or $b_1 = r\frac{s_y}{s_x}$

and

where $b_0 = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$ or $b_0 = \bar{y} - b_1\bar{x}$

Confidence interval for y-intercept $\beta_0$

$b_0 - E < \beta_0 < b_0 + E$

where $E = t_{\alpha/2}\, s_e \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum x^2 - \frac{(\sum x)^2}{n}}}$

Confidence interval for slope $\beta_1$

$b_1 - E < \beta_1 < b_1 + E$

where $E = t_{\alpha/2} \frac{s_e}{\sqrt{\sum x^2 - \frac{(\sum x)^2}{n}}}$

## Chapter 11

$$\chi^2 = \sum \frac{(O-E)^2}{E} \text{ where } E = \frac{(\text{row total})(\text{column total})}{\text{sample size}}$$

Tests of Independence $d.f. = (R-1)(C-1)$

Goodness of fit $d.f. = (\text{number of categories}) - 1$

## Chapter 12

## One Way ANOVA

$k$ = number of groups; $N$ = total sample size

$$SS_{TOT} = \sum x^2_{TOT} - \frac{(\sum x_{TOT})^2}{N}$$

$$SS_{BET} = \sum_{\text{all groups}} \left(\frac{(\sum x_i)^2}{n_i}\right) - \frac{(\sum x_{TOT})^2}{N}$$

$$SS_W = \sum_{\text{all groups}} \left(\sum x_i^2 - \frac{(\sum x_i)^2}{n_i}\right)$$

$$SS_{TOT} = SS_{BET} + SS_W$$

$MS_{BET} = \frac{SS_{BET}}{d.f._{BET}}$ where $d.f._{BET} = k-1$

$MS_W = \frac{SS_W}{d.f._W}$ where $d.f._W = N-k$

$F = \frac{MS_{BET}}{MS_W}$ where $d.f.$ numerator $= d.f._{BET} = k-1$

$d.f.$ denominator $= d.f._W = N-k$

### Two-Way ANOVA

$r$=number of rows; $c$=number of columns

Row factor $F$: $\frac{MS \text{ row factor}}{MS \text{ error}}$

Column factor $F$: $\frac{MS \text{ column factor}}{MS \text{ error}}$

Interaction $F$: $\frac{MS \text{ interaction}}{MS \text{ error}}$

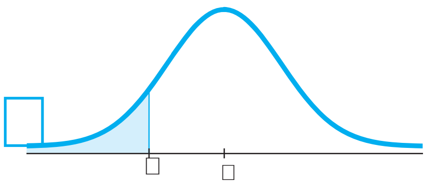
with degrees of freedom for

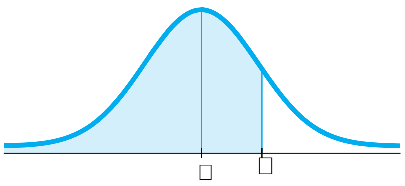
row factor = $r-1$

column factor = $c-1$

interaction = $(r-1)(c-1)$

error = $rc(n-1)$





critical z-valuesfor hypothesis testing

α = 0.10
c-level = 0.90

α = 0.05
c-level = 0.95

α = 0.01
c-level = 0.99

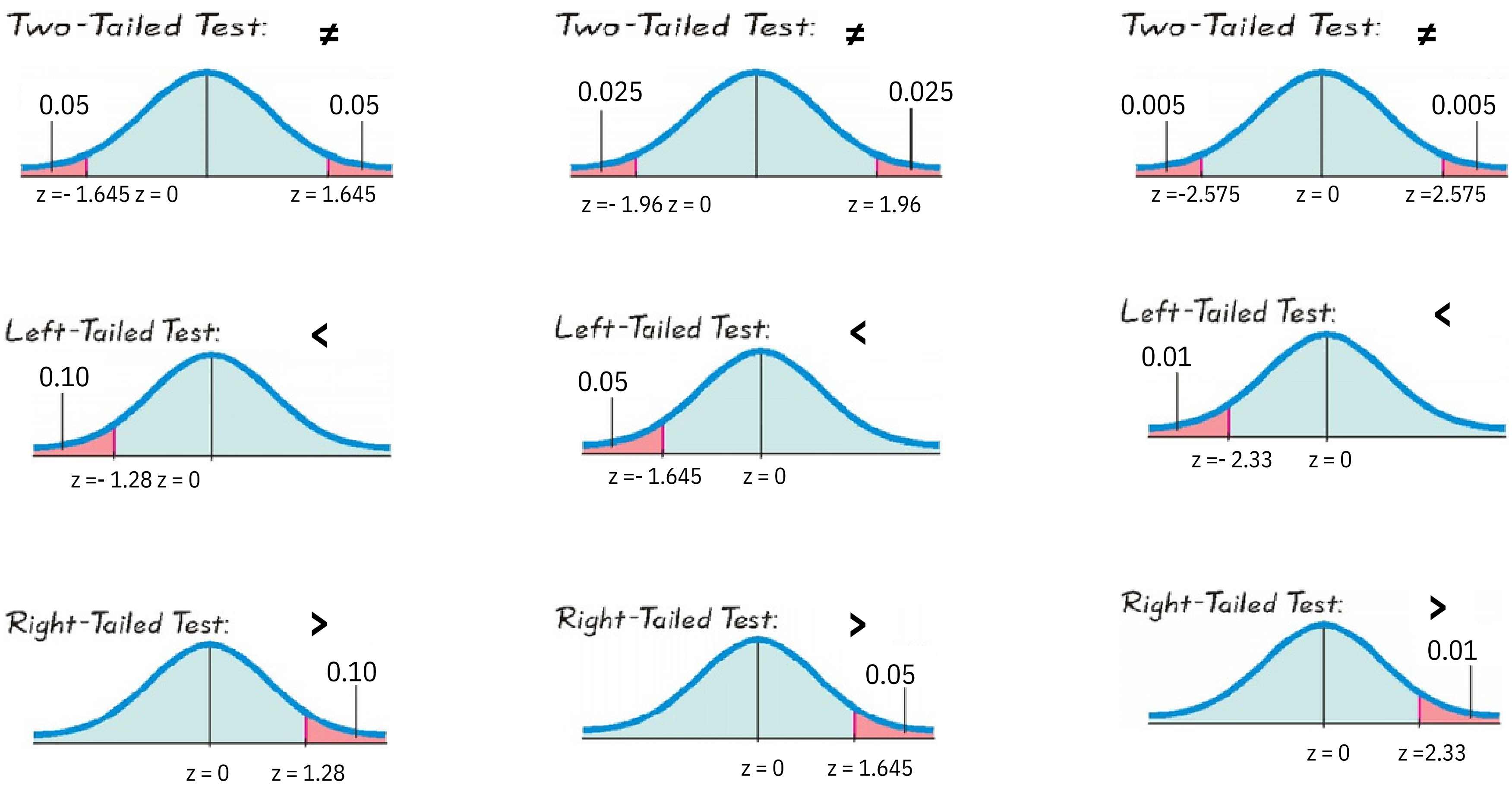



Figure 8.4

| confident that $s^2$ is within | the sample size $n$ should be at least | confident that $s$ is within | the sample size $n$ should be at least |
|---|---|---|---|
| 1% | 77,208 | 1% | 19,205 |
| 5% | 3,149 | 5% | 768 |
| 10% | 806 | 10% | 192 |
| 20% | 211 | 20% | 48 |
| 30% | 98 | 30% | 21 |
| 40% | 57 | 40% | 12 |
| 50% | 38 | 50% | 8 |
| To be 99% confident that $s^2$ is within | of the value of $\sigma^2$, the sample size $n$ should be at least | To be 99% confident that $s$ is within | of the value of $\sigma$, the sample size $n$ should be at least |
| 1% | 133,449 | 1% | 33,218 |
| 5% | 5,458 | 5% | 1,336 |
| 10% | 1,402 | 10% | 336 |
| 20% | 369 | 20% | 85 |
| 30% | 172 | 30% | 38 |
| 40% | 101 | 40% | 22 |
| 50% | 68 | 50% | 14 |

(table 7-2 from page 390, Triola 4th edition)

| | | |
|---|---|---|
| [illegible] | [illegible] | [illegible] |
| 10 | 0.632 | 0.765 |
| 11 | 0.602 | 0.735 |
| 12 | 0.576 | 0.708 |
| 13 | 0.553 | 0.684 |
| 14 | 0.532 | 0.661 |
| 15 | 0.514 | 0.641 |
| 16 | 0.497 | 0.623 |
| 17 | 0.482 | 0.606 |
| 18 | 0.468 | 0.590 |
| 19 | 0.456 | 0.575 |
| 20 | 0.444 | 0.561 |
| 25 | 0.396 | 0.505 |
| 30 | 0.361 | 0.463 |
| 35 | 0.335 | 0.430 |
| 40 | 0.312 | 0.402 |
| 45 | 0.294 | 0.378 |
| 50 | 0.279 | 0.361 |
| 60 | 0.254 | 0.330 |
| 70 | 0.236 | 0.305 |
| 80 | 0.220 | 0.286 |
| 90 | 0.207 | 0.269 |
| 100 | 0.196 | 0.256 |

NOTE: To test H0: $\rho = 0$ against H1: $\rho \neq 0$, reject H0 if the absolute value of r is greater than the critical value in the table.

Greek Alphabet

| Greek Letter | | Name | Equivalent | Sound When Spoken |
|---|---|---|---|---|
| Α | α | Alpha | A | al-fah |
| Β | β | Beta | B | bay-tah |
| Γ | γ | Gamma | G | gam-ah |
| Δ | δ | Delta | D | del-tah |
| Ε | ε | Epsilon | E | ep-si-lon |
| Ζ | ζ | Zeta | Z | zay-tah |
| Η | η | Eta | E | ay-tay |
| Θ | θ | Theta | Th | thay-tah |
| Ι | ι | Iota | I | eye-o-tah |
| Κ | κ | Kappa | K | cap-ah |
| Λ | λ | Lambda | L | lamb-dah |
| Μ | μ | Mu | M | mew |
| Ν | ν | Nu | N | new |
| Ξ | ξ | Xi | X | zzEye |
| Ο | ο | Omicron | O | om-ah-cron |
| Π | π | Pi | P | pie |
| Ρ | ρ | Rho | R | row |
| Σ | σ | Sigma | S | sig-ma |
| Τ | τ | Tau | T | tawh |
| Υ | υ | Upsilon | U | oop-si-lon |
| Φ | φ | Phi | Ph | figh or fie |
| Χ | χ | Chi | Ch | kigh |
| Ψ | ψ | Psi | Ps | sigh |
| Ω | ω | Omega | O | o-may-gah |